WHAT IS CLAIMED IS:

1. A system for processing an audio signal comprising:

means for dividing the audio signal into segments, each segment representing a portion of the audio signal occurring in one of a succession of time intervals;

means for detecting for each segment the presence of a fundamental frequency;

means responsive to the detecting means for determining the voicing probability for each segment by computing a ratio between voiced and unvoiced components of the audio signal, the determining means comprising:

means for windowing each segment of the audio signal;

means for computing the spectrum of the windowed segment;

means for computing correlation coefficients of each segment using at least the spectrum; and

means for comparing the correlation coefficients with a voicing threshold for each segment;

means for separating the signal in each segment into a voiced portion and an unvoiced portion on the basis of the voicing probability, wherein the voiced portion of the signal occupies the low end of the spectrum and the unvoiced portion of the signal occupies the high end of the spectrum for each segment; and

means for separately encoding the voiced portion and the unvoiced portion of the audio signal.

2. The system of Claim 1, wherein the audio signal is a speech signal and the means for determining the voicing probability further comprises means for refining the fundamental frequency of each segment using at least the spectrum of the windowed segment.

3. The system of Claim 1, wherein the means for encoding comprises means for computing LPC coefficients for a speech segment and means for transforming LPC coefficients into line spectral frequencies (LSF) coefficients corresponding to the LPC coefficients.

4. The system of Claim 1, wherein the means for computing the spectrum of the windowed segment comprises means for performing a Fast Fourier Transform (FFT) of the windowed segment.

5. The system of Claim 1, further comprising means for estimating the voicing threshold for each segment comprising:

means for dividing the spectrum into a plurality of non-linear bands, where the low bands of the spectrum have a higher resolution than the high bands of the spectrum;

means for evaluating at least one voice measurement for each of the plurality of bands, where the at least one voice measurement is the normalized correlation coefficients calculated in the frequency domain;

means for computing the low band energy of the spectrum;

means for computing an energy ratio between the energy of the high and low bands of the spectrum of a current segment and a previous segment; and

a multi-layer neural network classifier for receiving the normalized correlation coefficients of the low bands, the low band energy and the energy ratio.

6. The system of Claim 1, further comprising means for spectrally estimating the audio signal comprising:

means for calculating a complex spectrum for each segment by using a window based on the fundamental frequency;

means for spectrally modeling each segment using at least the complex spectrum, the fundamental frequency, and the voicing probability to obtain line spectral frequencies (LSF) coefficients and a signal gain of each segment.

7. The system of Claim 6, wherein the means for calculating the complex spectrum comprises means for applying a Fast Fourier Transform to the windowed segment.

8. A system for processing an audio signal comprising:

means for dividing the signal into segments, each segment representing a portion of the audio signal in one of a succession of time intervals;

means for detecting for each segment the presence of a fundamental frequency;

means responsive to the detecting means for determining the voicing probability for each segment by computing a ratio between voiced and unvoiced components of the audio signal;

means for calculating a complex spectrum for each segment by using a window based on the fundamental frequency;

means for spectrally modeling each segment using at least the complex spectrum, the fundamental frequency, and the voicing probability to obtain line spectral frequencies (LSF) coefficients and a signal gain of each segment;

means for separating the signal in each segment into a voiced portion and an unvoiced portion on the basis of the voicing probability, wherein the voiced portion of the signal occupies the low end of the spectrum and the unvoiced portion of the signal occupies the high end of the spectrum for each segment; and

means for separately encoding the voiced portion and the unvoiced portion of the audio signal.

9. The system of Claim 8, wherein the audio signal is a speech signal and the means for determining the voicing probability comprises means for refining the fundamental frequency of each segment using at least the spectrum of the windowed segment.

10. The system of Claim 8, wherein the means for encoding comprises means for computing LPC coefficients for a speech segment and means for transforming LPC coefficients into line spectral frequencies (LSF) coefficients corresponding to the LPC coefficients.

11. The system of Claim 8, wherein the means for computing the spectrum of the windowed segment comprises means for performing a Fast Fourier Transform (FFT) of the windowed segment.

12. The system of Claim 8, wherein the means for determining the voicing probability comprises:

means for windowing each segment of the input signal;

means for computing the spectrum of the windowed segment;

means for computing correlation coefficients of each segment using at least the spectrum; and

means for comparing the correlation coefficients with a voicing threshold for each segment.

13. The system of Claim 12, further comprising means for estimating the voicing threshold for each segment comprising:

means for dividing the spectrum into a plurality of non-linear bands, where the low bands of the spectrum have a higher resolution than the high bands of the spectrum;

means for evaluating at least one voice measurement for each of the plurality of bands, where the at least one voice measurement is the normalized correlation coefficients calculated in the frequency domain;

means for computing the low band energy of the spectrum;

means for computing an energy ratio between the energy of the high and low bands of the spectrum of a current segment and a previous segment; and

a multi-layer neural network classifier for receiving the normalized correlation

coefficients of the low bands, the low band energy and the energy ratio.

14. The system of Claim 8, wherein the means for calculating the complex spectrum comprises means for applying a Fast Fourier Transform to the windowed segment.

15. A system for processing an audio signal having a number of frames, the system comprising:

an encoder comprising:

first means for determining for each frame a ratio between voiced and unvoiced components of the audio signal on the basis of the fundamental frequency of each frame, the ratio being defined as a voicing probability, the means for determining the voicing probability comprising:

means for windowing each frame of the input signal;

means for computing the spectrum of the windowed frame;

means for computing correlation coefficients of each frame using at least the spectrum; and

means for comparing the correlation coefficients with a voicing threshold for each segment;

second means for determining at least a pitch period, a mid-frame pitch period, and/or a mid-frame voicing probability of the audio signal; and

means for quantizing at least the pitch period, the voicing probability, the mid-frame pitch period, and/or the mid-frame voicing probability.

16. The system of Claim 15, further comprising a decoder comprising:

means for unquantizing at least the pitch period, the voicing probability, the mid-frame pitch period, and/or the mid-frame voicing probability and providing at least one output; and

means for analyzing the at least one output to produce a synthetic speech signal corresponding to the input audio signal.

17. The system of Claim 15, further comprising means for estimating the voicing threshold for each segment comprising:

means for dividing the spectrum into a plurality of non-linear bands, where the low bands of the spectrum have a higher resolution than the high bands of the spectrum;

means for evaluating at least one voice measurement for each of the plurality of bands, where the at least one voice measurement is the normalized correlation coefficients calculated in the frequency domain;

means for computing the low band energy of the spectrum;

means for computing an energy ratio between the energy of the high and low bands of the spectrum of a current segment and a previous segment; and

means for receiving the normalized correlation coefficients of the low bands, the low band energy and the energy ratio.

18. The system of Claim 17, wherein the means for receiving is a multi-layer neural network classifier.

19. The system of Claim 18, wherein the voicing probability is zero if an output from the means for receiving is less than a predetermined threshold for a predetermined number of frames.

20. The system of Claim 15, wherein further comprising means for high-pass filtering the audio signal and buffering the audio signal into the number of frames.

21. The system of Claim 15, wherein the encoder further comprises spectral estimation means for computing an estimate of the power spectrum of the audio signal using a pitch adaptive window.

spectral magnitude envelope;

means for calculating sine-wave phases using at least the calculated frequencies; and

means for calculating a sum of sinusoids using at least the calculated frequencies and amplitudes and the sine-wave phases to produce the time-domain signal.

26. The system of Claim 15, further comprising:

means for calculating a complex spectrum for each segment by using a window based on the fundamental frequency; and

means for spectrally modeling each segment using at least the complex spectrum, the fundamental frequency, and the voicing probability to obtain line spectral frequencies (LSF) coefficients and a signal gain of each segment.

27. The system of Claim 26, wherein the means for calculating the complex spectrum comprises means for applying a Fast Fourier Transform to the windowed segment.

28. A system for processing an audio signal having a number of frames, the system comprising:

an encoder comprising:

means for determining for each frame a ratio between voiced and unvoiced components of the audio signal on the basis of the fundamental frequency of each frame, the ratio being defined as a voicing probability;

means for calculating a complex spectrum for each segment by using a window based on the fundamental frequency;

means for spectrally modeling each segment using at least the complex spectrum, the fundamental frequency, and the voicing probability to obtain line spectral frequencies (LSF) coefficients and a signal gain of each segment;

means for determining at least a pitch period, a mid-frame pitch

22. The system of Claim 21, wherein the length of the pitch adaptive window is based on the fundamental frequency of the audio signal.

23. The system of Claim 16, wherein the means for unquantizing comprises:

means for producing a spectral magnitude envelope and a minimum phase envelope using at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability;

means for interpolating and outputting the spectral magnitude envelope and the minimum phase envelope to the means for analyzing;

means for estimating the signal-to-noise ratio of the audio signal using the at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability; and

means for generating at least one control parameter using at least the signal-to-noise ratio and for outputting the at least one control parameter to the means for analyzing.

24. The system of Claim 16, wherein the means for analyzing comprises:

first means for processing the at least one output to produce a time-domain signal; and

second means for processing the time-domain signal to produce the synthetic speech signal corresponding to the audio signal.

25. The system of Claim 24, wherein the first means for processing the at least one output to produce the time-domain signal comprises:

means for filtering a spectral magnitude envelope, wherein the spectral magnitude envelope is outputted by the means for unquantizing;

means for calculating frequencies and amplitudes using at least the filtered

period, and/or a mid-frame voicing probability of the audio signal; and

means for quantizing at least the pitch period, the voicing probability, the mid-frame pitch period, and/or the mid-frame voicing probability.

29. The system of Claim 28, further comprising a decoder comprising:

means for unquantizing at least the pitch period, the voicing probability, the mid-frame pitch period, and/or the mid-frame voicing probability and providing at least one output; and

means for analyzing the at least one output to produce a synthetic speech signal corresponding to the input audio signal.

30. The system of Claim 28, further comprising means for estimating the voicing threshold for each segment comprising:

means for dividing the spectrum into a plurality of non-linear bands, where the low bands of the spectrum have a higher resolution than the high bands of the spectrum;

means for evaluating at least one voice measurement for each of the plurality of bands, where the at least one voice measurement is the normalized correlation coefficients calculated in the frequency domain;

means for computing the low band energy of the spectrum;

means for computing an energy ratio between the energy of the high and low bands of the spectrum of a current segment and a previous segment; and

means for receiving the normalized correlation coefficients of the low bands, the low band energy and the energy ratio.

31. The system of Claim 30, wherein the means for receiving is a multi-layer neural network classifier.

32. The system of Claim 31, wherein the voicing probability is zero if an output from the means for receiving is less than a predetermined threshold for a predetermined number of frames.

33. The system of Claim 28, further comprising means for high-pass filtering the audio signal and buffering the audio signal into the number of frames.

34. The system of Claim 28, wherein the encoder further comprises spectral estimation means for computing an estimate of the power spectrum of the audio signal using a pitch adaptive window.

35. The system of Claim 34, wherein the length of the pitch adaptive window is based on the fundamental frequency of the audio signal.

36. The system of Claim 29, wherein the means for unquantizing comprises:

means for producing a spectral magnitude envelope and a minimum phase envelope using at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability;

means for interpolating and outputting the spectral magnitude envelope and the minimum phase envelope to the means for analyzing;

means for estimating the signal-to-noise ratio of the audio signal using the at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability; and

means for generating at least one control parameter using at least the signal-to-noise ratio and for outputting the at least one control parameter to the means for analyzing.

37. The system of Claim 29, wherein the means for analyzing comprises:

first means for processing the at least one output to produce a time-domain signal; and

second means for processing the time-domain signal to produce the synthetic speech signal corresponding to the audio signal.

38. The system of Claim 37, wherein the first means for processing the at least one output to produce the time-domain signal comprises:

means for filtering a spectral magnitude envelope, wherein the spectral magnitude envelope is outputted by the means for unquantizing;

means for calculating frequencies and amplitudes using at least the filtered spectral magnitude envelope;

means for calculating sine-wave phases using at least the calculated frequencies; and

means for calculating a sum of sinusoids using at least the calculated frequencies and amplitudes and the sine-wave phases to produce the time-domain signal.

39. The system of Claim 28, wherein the means for determining the voicing probability comprises:

means for windowing each frame of the input signal;

means for computing the spectrum of the windowed frame;

means for computing correlation coefficients of each frame using at least the spectrum; and

means for comparing the correlation coefficients with a voicing threshold for each segment.

40. The system of Claim 28, wherein the means for calculating the complex spectrum comprises means for applying a Fast Fourier Transform to the windowed segment.

41. A system for processing an audio signal having a number of frames, the system comprising:

a decoder comprising:

means for unquantizing at least a pitch period, a voicing probability, a mid-frame pitch period, and/or a mid-frame voicing probability of the audio signal and providing at least one output, where the means for unquantizing comprises means for generating at least one control parameter using at least the signal-to-noise ratio computed using a gain and the voicing probability of the audio signal; and

means for analyzing the at least one output, including the at least one control parameter, to produce a synthetic speech signal corresponding to the input audio signal.

42. The system of Claim 41, wherein the means for unquantizing comprises:

means for producing a spectral magnitude envelope and a minimum phase envelope using at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability;

means for interpolating and outputting the spectral magnitude envelope and the minimum phase envelope to the means for analyzing; and

means for estimating the signal-to-noise ratio of the audio signal using the at least the unquantized pitch period, the unquantized voicing probability, the unquantized mid-frame pitch period, and/or the unquantized mid-frame voicing probability and outputting the signal-to-noise ratio to the means for generating at least one control parameter.

43. The system of Claim 41, wherein the means for analyzing comprises:

first means for processing the at least one output to produce a time-domain signal; and

second means for processing the time-domain signal to produce the synthetic

speech signal corresponding to the audio signal.

44. The system of Claim 43, wherein the first means for processing the at least one output to produce the time-domain signal comprises:

means for filtering a spectral magnitude envelope, wherein the spectral magnitude envelope is outputted by the means for unquantizing;

means for calculating frequencies and amplitudes using at least the filtered spectral magnitude envelope;

means for calculating sine-wave phases using at least the calculated frequencies; and

means for calculating a sum of sinusoids using at least the calculated frequencies and amplitudes and the sine-wave phases to produce the time-domain signal.